DOCUMENT RESUME

ED 201 672 TM 810 268

AUTHOR Hansen, Joe B.
TITLE Uses of Title I Evaluation Data: Implications for Technical Assistance.
PUB DATE Apr 81
NOTE 17p.: Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Decision Making: Elementary Secondary Education: *Research Utilization: *Technical Assistance: *Use Studies
IDENTIFIERS Evaluation Problems: *Evaluation Utilization: Technical Assistance Centers: *Title I Evaluation and Reporting System

ABSTRACT

The basic position taken in this paper is that the data collected under the Title I evaluation and reporting system can be used more effectively than they have been traditionally in making decisions about Title I programs, and that other types of data can and should be used in shaping and improving Title I programs. This paper combines information derived from interviews conducted with state Title I coordinators and evaluators in nine of the 14 states served by the Title I Evaluation, Technical Assistance Centers for regions 8, 9, and 10 and a review of the literature on evaluation utilization, to provide guidance for planning technical assistance services. To provide an analytical framework for the literature review, four sets of hypothetical factors influencing evaluation were identified: technical-methodological, personal-role related, communication related, and policical-organization. These four factors also provided categories for classifying problems in evaluation use for the interviews. State Education Agency respondents suggested three areas of technical assistance as having potential for improving evaluation use: data interpretation, implementation evaluation, and longitudinal (sustained effects) evaluation. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED201672

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

USES OF TITLE I EVALUATION DATA: IMPLICATIONS FOR TECHNICAL ASSISTANCE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. B. Hansen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Joe B. Hansen

Northwest Regional Educational Laboratory

Paper presented at the annual meeting of the American Educational Research Association. Los Angeles, CA: April 1981.

TM 810 268

Acknowledgements

The contribution of the following individuals are gratefully acknowledged:

Dr. Barbara Williams for assistance in conducting interviews.

Dr. Gary Estes for his review and suggestions for improvement.

Dr. Stephen Murray for his review and suggestions.

Implications for Technical Assistance

Joe B. Hansen
Northwest Regional Educational Laboratory
Portland, Oregon

Paper presented at the annual meeting of the American Educational Research Association. Los Angeles, CA: April 1981.

Introduction

Over the last several years a rapidly growing body of evaluation literature has emerged on the topic of "Evaluation Utilization" or simply "Evaluation Use". This body of literature has grown out of a concern over the apparent lack of use of the evaluation studies that are conducted by the hundreds annually. While there has been an increasing emphasis on evaluation, the literature is rife with evidence that the use of evaluation results in making program decisions is far below that which might reasonably be expected. That this is a continuing problem can be illustrated by two excerpts from the literature. Weiss (1972) wrote--

> "Evaluation research is meant for immediate and direct use in improving the quality of social programming, yet a review of evaluation experience suggests that evaluation results have not exerted significant influence on program decisions."

More recently Wise, reporting at AERA in 1978 stated,

> "Evaluation has for some time been defined as a process of gathering and analyzing information for decision makers. The problem seems to be that the decision makers do not read or use the information provided, or use it towards purposes that the evaluator does not feel are legitimate."

While these quotations voice strong pessimism other evaluation researchers have stated that evaluations are being used (Alkin, Daillak and White, 1979; Patton, 1978). This apparent disagreement reflects a problem in defining utilization and selecting criteria to determine the influence of evaluation on decision makers. An often quoted definition of evaluation is "Evaluation is the process of delineating, obtaining and providing useful information for judging decision alternatives" (Stufflebeam, Foley, Gephart, Guba, Hammond, Merriman and Provus 1979, pg. 36) Alkin, et al. (1979) proposed a more extensive definition of utilization with the hope of suggesting a workable framework for future research on evaluation utilization. According to Alkin et al. (1979), the utilization of evaluation information occurs when users consider the information as providing some influence in making a decision about the future content, practice, appearance or existence of the program. Each of four factors, the users, amount of influence, type of decision and future direction is further described by listing common alternatives the authors found in case studies.

Any attempt to improve evaluation use should begin with a workable definition of evaluation use. As it is used in this paper evaluation use involves elements of both above definitions. Whereas the Stufflebeam, et al. (1971) definition emphasizes the technical features of evaluation and requires a certain level of data quality be achieved before utilization should occur, Alkin et al. (1979) focuses primarily on utilization as it occurs in the context of a social program. The basic position taken in this paper is that the data collected under the Title I evaluation and reporting system can be used more effectively than they have been traditionally in making decisions about Title I programs and that other types of data can and should be used in shaping and improving Title I programs.

This paper combines information derived from interviews conducted with state Title I coordinators and evaluators in nine of the 14 states served by the Title I Evaluation, Technical Assistance Centers for regions 8, 9, and 10 and a review of the literature on evaluation utilization (Hansen, Martin, and Oxford, 1980), to provide guidance for planning technical assistance services.

The Literature Review

The literature review addressed two questions:

1. What are the factors limiting the usefulness of evaluation?

2. How can evaluation be made more useful?

To provide an analytical framework for the literature review, four sets of hypothetical factors influencing evaluation use were identified: Technical-methodological, personal-role related, communication related and political-organizational. These four factors provided categories for classifying problems in evaluation use. Each factor is discussed briefly below.

Technical-methodological factors

Technical-methodological factors are the most familiar and commonly identified obstacles to evaluation use. They involve the mechanics of the evaluation and how deficiencies or oversites within the evaluation process affect data quality and use. These factors include considerations of design, instrumentation, procedures and data analysis techniques. Reliability and validity data are the aspects of the data most affected by these factors. A partial guide to the influence of technical factors in evaluation includes Tallmadge and Wood (1976), Weiss (1972), Campbell and Stanley (1963) and Rossi (1972).

Personal-role-related-factors

Various personal and role-related factors may impede the efficient use of evaluation results by decision makers. According to Gurel (1975) the most important barriers to utilization of evaluation results are not technical but interpersonal. Other researchers (Longood, Simmel, 1972) have identified personality factors as obstacles to evaluation use while the differing roles of managers and evaluators have been identified by Agarwala-Rogers (1977) Carter (1973) and Weiss (1975) as posing obstacles to evaluation use. Personality or value preference factors are also important influences on how information is used. The information needs of decision makers vary as a function of role with managers focusing more on process information and evaluators focusing on product information. This category of factors can include psycho-social differences as well as role characteristics.

Communication Factors

Communications problems that may interfere with the use of evaluation results include: credibility of the information source (Alkin et. al., 1979) types of information, timeliness of the information (Weiss, 1972; Agarwala-Rogers, 1977), influence of the information on decision making (Nielsen, 1975; Hansen, 1978), communication linkages (Halpert, 1969) and evaluator advocacy (Weiss, 1972).

Political-organizational factors

The importance of political and organizational factors in evaluation use has been recognized by various writers (Patton, 1978; Morris & Fitzgibbon, 1978; Weiss, 1975). School administrators are responsible to taxpayers and are thus vulnerable to pressures exerted on them by special interest groups and other external sources of influence. Nielsen (1975) and Hansen (1978) have noted that useful evaluation studies must be based on an understanding of the decision making process and the political context in which it operates.

While the literature review was quite helpful in developing a very general contour map of evaluation use problems and strategies for overcoming them, it bore limited implications for TAC services to improve evaluation use.

During the summer of 1980 I initiated a series of discussions with TAC state contact persons in the states served by the Title I TAC for regions 8, 9 and 10 on the issue of evaluation use. These discussions were originally intended to provide us with information for targeting TAC services during the 1980-81 service year. However, scheduling and other problems interfered to some extent and the discussions were extended through February of 1981. These discussions or informal interviews sought the perceptions of TAC contact persons on two issues parallel to those examined in the literature review: 1) what are the major obstacles

facings LEAs and SEAs in making evaluations useful? 2) how can the services of the title I TACs help to improve evaluation use? A total of nine interviews were conducted, each with a different state contact person. In the remainder of this presentation the results of those interviews are examined in terms of the four factors derived from the literature and implications for TAC services to assist in the improvement of evaluation use are examined.

Interview Results

The results of the interviews are summarized in Table 1. Each response was characterized in brief terms from notes taken during the discussion. Responses were then coded according to the four factor framework of the literature review (T=technical, C=communications, P=political and R=role-related or personal). TAC service needs identified by the respondents have been listed in brief form in column four.

Obstacles to evaluation use

As shown in Table 1, six respondents identified technical factors as posing obstacles to evaluation use. Since in many cases more than one obstacle was suggested, the number of obstacles identified is greater than the number of respondents, with nine total responses in the technical category. Five obstacles were identified as communications related and three were identified as role-related. Within the technical category lack of reliability and or validity in the data was the most frequently cited obstacle to effective use, with three of the respondents referring to this issue in some form. Other technical obstacles included lack of skill in evaluation procedures and practical or real world environmental constraints on evaluation designs. In the communications category, the lack of synchronization of the reporting or planning cycle was identified by two respondents. Only two of the obstacles reported were clearly political; one respondent cited the suppression of negative results to protect favorite programs, while another cited the use of negative results to shake up the programs. A lack of longitudinal perspective at the classroom level and lack of involvement of key decision makers in the evaluation process, were identified as role-related problems.

Although most obstacles to evaluation use identified were technical in nature they represented a wide variety of problems rather than any clearly defined core of common concerns. To some extent this reflects differences in conditions within the states, and to some extent it reflects differences in the interests, skills and philosophies of the interviewees. For example, in one state where data quality is of major importance, a considerable effort has been mounted by the state to improve the quality of data through centralized data collection and the use of a computer to reduce errors in computation. This state was one of the first to implement the TIERS computer program and has kept abreast of all changes and improvements in the use of computers for TIERS data analysis.

Alternative definitions of "evaluation use" were seen as important in a state wherein the contact person was a professional evaluator. This person's interest in pursuing research questions at the local level was also reflected in obstacles cited.

In summary, most of the obstacles to evaluation use cited by the state contacts were technical problems relating to model implementation, measurement and data quality issues. These responses varied, however, revealing no clear common concern other than the apparent conviction that quality is a necessary pre-condition for effective use.

TAC Service Needs

The second general question asked of the state contacts was,"How can Title I TAC services help to improve evaluation use?" These responses are summarized in Table 1, column 4.

Most suggestions for targeting TAC services focused on technical strategies, regardless of the nature of the obstacle identified. For example, the use of negative results to "shake up" a program is a political use of the data, yet the proposed TAC service strategy is one that focuses on technical concerns regarding the use of data.

Sustained effects and implementation evaluation were each identified as T.A. need areas by three states, with one SEA contact mentioning both of these. More training on data interpretation and use was also identified as a need by three respondents. Assistance with local studies and a refocusing of evaluation at the building level were also seen as helpful.

The SEA contacts tended, as a group, to perceive a need for assistance on technical problems as an appropriate means to improve evaluation use, regardless of the nature of the obstacle they identified. The most commonly cited TAC service needs were for assistance in sustained effects measurement, data interpretation and implementation evaluation.

It may be as important to note certain omissions from the responses as it is to note the responses themselves. For example, in an attempt to become more cost efficient and to create a legacy of instructional and reference materials which will outlast the TACs themselves, the Title I TACs nationally are engaged in a coordinated effort to produce high quality instructional materials aimed at specific evaluation problems. While this effort has been under way for more than a year, none of the state contact persons mentioned instructional materials either in their problem or their solution statements.

It appears that after four years, there is still some frustration with perceived deficiencies in the TIERS system, most notably, its interpretability, and with the inherent usefulness of NCEs and aggregated achievement data to program planners and implementers. In general, TAC services seem to be perceived as useful and as having at least potential value in solving some of the problems related to evaluation use.

Observations and Implications

The purpose of the literature review was to produce a document (Hansen, et al., 1980) that identified the major factors influencing evaluation use and provided suggestions for evaluators and other practitioners on ways to improve the use of evaluation studies. It was not conceived as a planning document for TAC services but as an evaluator's reference. The interviews, on the other hand, were but one part of the routine information gathering that takes place in the NWREL TACs in an attempt to sense T.A. needs and target our resources most effectively. Other more detailed forms of needs assessment also take place, including surveys of workshop participants, follow-up telephone calls and other approaches. The results of either of these two activities, taken independently provide, at best, weak indicators of T.A. needs. Their joint results however, may have some general implications for providing T.A. services in Title I evaluation over the next one and one-half years.

The SEA contacts suggested three areas of technical assistance as having potential for improving evalution use: data interpretation, implementation evaluation and longitudinal (sustained effects) evaluation. Of these, only "implementation evaluation" was not a major source of concern in the literature. This is probably due to the fact that although implementation research has been with us for some time, this concept has, with notable exceptions such as Stufflebeam et al. (1971), received limited emphasis in evaluation. The Title I TACs have recently been at the forefront in exploring this area and developing technical assistance strategies to support it.

I had anticipated a stronger tendency for the SEA contacts to identify implementation problems and to suggest T.A. needs in more "qualitative" areas such as observational and process methodology, implying perhaps, a definite shift in TAC service, away from technical details toward broader issues. This failed to emerge. As a result, a "shift" in service seems less appropriate than a gradual transition, in which the approach to T.A. services remains balanced. Since the advent of TIERS, the services provided by TACs have undergone a continual gradual transition from awareness level training on the TIERS requirements to skill development in measurement and research techniques. An emergent trend has been toward a more generic "evaluation improvement" orientation. In many cases, however, we have found the remediation of technical skills deficits to be a necessary pre-condition for improvement of evaluation quality. This does not imply that effective use of evaluation depends primarily on quality of data in traditional terms, i.e., validity, reliability, etc.

The lack of emphasis by SEA respondents on the qualitative areas of evaluation as areas of need for T.A. could be explained in two ways: (1) a built-in bias toward more technical--measurement oriented--issues may have resulted from the fact that the interviews were conducted by the Technical Assistance Center; (2) a tendency to view evaluation from a somewhat restrictive, summative--measurement oriented--perspective. In either case there may be implications for working with clients to broaden their view of evaluation and potential areas for TAC services. This has been an active part of the TAC agenda over the past two to three years.

Consistent with the notion of broadening the concept of evaluation is the idea of extending the concept of "use" beyond a simple dichotomy, as Alkin et al. (1979) have done. As illustrated in figure 1, "evaluation use" may vary in terms of the extent to which evaluation results are used in program decision making.

Figure 1

Degrees of Use of Evaluation Results
In Decision Making

Program Awareness or Public Relations decision process	Some influence on or consideration in decision making	Major basis for decision making
1) Report results to PAC or school board	1) Comparison of rotating program serving many kids--short-term vs. pull-out serving a few kids--long term	1) Decision to seek JDRP approval/funding
2) Newspaper reports	2) Sustaining effects type of study	2) Year-round school programs or other Title IV-A type program

It may be possible, by viewing evaluation use as a more continuous variable in this fashion, to tease out relationships among evaluation consumers, levels of use and emphasis for T.A. For example, if one accepts the notion that evaluation problems and uses are role related then it follows that evaluation T.A. needs should be role related. Referring to figure 1 again, it seems logical to focus training for PACs or other groups for whom increasing awareness is the main purpose, on communications related topics such as interpretation of results from displays. Program coordinators and other building level administrators may need training which strikes a balance between technical and communicative emphases, whereas evaluators might require more of a focus on technical issues. I don't mean to imply that these relationships are clear cut and specific but it may be helpful to us in planning T.A. programs to keep these concepts in mind.

I also expected the interviews to reveal a greater product orientation than they did, given the large scale effort the TACs have mounted nationally to develop quality instructional and reference materials. Most of the materials we have developed, however, are directed at the LEA. Therefore, the SEA contacts may not have perceived their usefulness as strongly as the LEAs do. This, at least, implies a need for a follow up with LEA clientele to ascertain the relative value of materials and services.

Epilogue

Since the efforts discussed here were initiated, numerous political events have occurred, exerting their influence on the LEAs, SEAs, TACs and others concerned with the quality of education. The advent of block grants may have far reaching and dramatic implications for education as a whole, evaluation as a discipline and technical assistance as a means to effect change. Whatever happens, it is likely that there will be a continuing and perhaps even increased need for T.A. as LEAs and SEAs attempt to adjust to the new requirements. There will be new demands for planning, monitoring and accountability functions as educators try to maintain educational quality with reduced levels of funding. Therefore, it seems that we in the TACs must concern ourselves with two types of needs as we plan for the future. We must continue to address the TIERS implementation requirements in the short run while at the same time looking ahead and beginning to prepare for an educational environment in which accountability will likely emerge as the dominant trend in evaluation, much as it was a decade ago.

REFERENCES

Agarwala-Rogers, R. Why is evaluation research not utilized? In M. Guttentag (Ed.), Evaluation studies review annual (Vol.2.) Beverly Hills: Sage, 1977.

Alkin, M.C., Daillak, R.D., and White, P. Using evaluation: Does evaluation make a difference? Beverly Hills: Sage, 1979.

Campbell, D.T., and Stanley, J.C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1962.

Carter, R.K. Client's resistance to negative findings, In E.R. House (Ed.), School evaluation: The politics and process. Berkeley: McCutchan, 1973.

Gurel, L. The human side of evaluating human services programs: Problems and prospects. In M. Guttentag and E.L. Struening (Eds.), Handbook of evaluation research (Vol 2). Beverly Hills: Sage, 1975.

Halpert, H.P. Communications as a basic tool in promoting utilization of research findings. In H.C. Schulberg, A. Sheldon, and F. Baker (Eds.), Program evaluation in the health fields. New York: Behavioral Publications, 1969.

Hansen, J.B. The formative-summative dichotomy and the role of description in program evaluation. Evaluation News, August 1978 (no. 6).

Hansen, J.B., and Oxford, R.L. Effects of the Title I evaluation and reporting system: TAC perspective. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1979.

Hansen, J.B., Martin, Joy M., and Oxford, R.L. The use of evaluation: An analytical review. Portland, OR: Northwest Regional Educational Laboratory, July, 1980.

Longood, R., and Simmel, A. Organizational resistance to innovation suggested by research. In C.H. Weiss (Ed.), Evaluating action programs: Readings in social action and education. Boston: Allyn and Bacon, 1972.

Nielsen, V.G. Why evaluation does not improve program effectiveness. Policy Studies Journal, 1975, 3 (4), 385-390.

Rossi, P.H. Boobytraps and pitfalls in the evaluation of social action programs. In C.H. Weiss (Ed.), Evaluating action programs: Readings in social action and education. Boston: Allyn and Bacon, 1972.

Stufflebeam, D.L., Foley, W.J., Gephart, W.J., Guba, E.G., Hammond, R.L., Merriman, H.O., and Provus, M.M. Educational evaluation and decision-making in education. Itasca, Ill.: Peacock, 1971.

Tallmadge, G.K., and Wood, C.T. User's guide (ESEA Title I Evaluation and Reporting System). Mountain View, CA: RMC Research Corporation, 1976.

Weiss, C.H. Utilization of evaluation: Toward comparative study. In C.H. Weiss (Ed.), Evaluating action programs: Readings in social action and education. Boston: Allyn and Bacon, 1972d. (Reprint of paper presented at the meeting of the American Sociological Association, Florida, 1966.)

Weiss, C.H. Evaluation research in the political context. In M. Guttentag and E.L. Struening (Eds.), Handbook of evaluation research (Vol. 1). Beverly Hills: Sage, 1975.

Wise, R. What we know about the decision-maker in decision settings. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

TABLE 1
Summary of
Interview Results

State	Major Obstacle	Category	TAC Service Need	Category
1	o Interpreting results (NCEs)	T	o Implementation evaluation training and materials to stimulate use	T
2	o Writing good objectives	C		
	o Data quality - Data are useful only to the extent they are reliable and valid	T	o T.A. on alternate solution to "regression effect"	T
			o Help LEAs learn how to ask useful questions in response to sustained effect requirements	T
3	o Two definitions of "use": a) Applied--focus on use in grades within school - Obstacle--improper model implementation, results=invalid data, invalid decisions b) Research--provide information for defining variables in quasi-experimental studies of program effectiveness and characteristics - constraints to doing experimental studies in the field, e.g., random assignment, control, etc.	T	o TAC help in making sure models are implemented correctly through LEA-TAC communication	T
			o Help LEAs learn of dangers of interpreting invalid data	T
			o Help LEAs identify valid designs for answering local questions	T
4	o Use of negative results to shake up programs. Not enough use of positive results	P	o Workshops on interpretation of results. Includes LEA training on selective use of results for pinpointing problems.	T
5	o Unreliability of data (quality)	T	o Little can be done to make TIERS data useful. TIERS is not organized for use at LEA level	R
	o Report cycle timing	C		
	o Confounding of results across programs, projects, classrooms	T	o Help LEA learn to use longitudinal data	T, C & R
			o Continue work to improve testing and measurement practices	T

State	Major Obstacle	Category	TAC Service Need	Category
6	o Emphasis on models and testing detracts from other possible kinds of useful data	T	o More training on sustained gain and implementation evaluation	T
	o Testing - planning cycle timing is off	C	o Focus more on building level, e.g., teachers and principals	C,R
	o Lack of longitudinal perspective at classroom level	R		
	o Target audience is administrator, should be teachers	R		
7	o Political - "lip service to integrity" If evaluation data do not favor or support programs favored by Supt. or other key administrators they are suppressed or ignored.	P	o Include district evaluators in TAC training as well as Title-I people. Promote coordination between district Title-I and evaluation and testing units.	P,C
	o Lack of Skill - LEAs show confusion; inability to identify information needs, users and methods for obtaining and using data.	T		
	Communications - an important skill area evaluators lack	C		
8	o Lack of Skill in use of "hard data" such as test score data for program decisions. Instruments in use are often inappropriate.	T	o LEA level training in what to expect from and how to use data	T
9	o Lack of Meaning of data being collected NCE gain lacks "intrinsic meaning" Doesn't help answer "what went wrong?"	C,T	o Program Improvement Evaluation - SEA & LEA & TAC should concentrate efforts on program improvement and implementation studies	T
	o Lack of Decision Maker Involvement in defining data needs and gathering data	R		

Obstacles

T=9, C=5, P=2, R=3

Needs

T=11, C=3, P=1, R=3